

7-23-2021


Linked Open Data, Wikidata-China, RDA and BIBFRAME: Status Report and Case Study

Sai Deng

University of Central Florida, sai.deng@ucf.edu

Amanda Xu

Lihong Zhu

 Part of the [Archival Science Commons](#), [Cataloging and Metadata Commons](#), [Collection Development and Management Commons](#), and the [Scholarly Publishing Commons](#)
Find similar works at: <https://stars.library.ucf.edu/ucfscholar>
University of Central Florida Libraries <http://library.ucf.edu>

This Conference Presentation is brought to you for free and open access by STARS. It has been accepted for inclusion in Faculty Scholarship and Creative Works by an authorized administrator of STARS. For more information, please contact STARS@ucf.edu.

Original Citation

Xu, A., Zhu, L.H. & Deng, S. (2021). Linked open data, Wikidata-China, RDA and BIBFRAME: Status report and case study. 2021 LD4 Conference on Linked Data. Online. July 23, 2021. Retrieved at: <https://ld42021.sched.com/list/descriptions/>

Linked Open Data, Wikidata-China, RDA and BIBFRAME: Status Report and Case Study

Amanda Z. Xu, National Agricultural Library
Lihong Zhu, Washington State University
Sai Deng, University of Central Florida






LD4 Conference, July 23, 2021



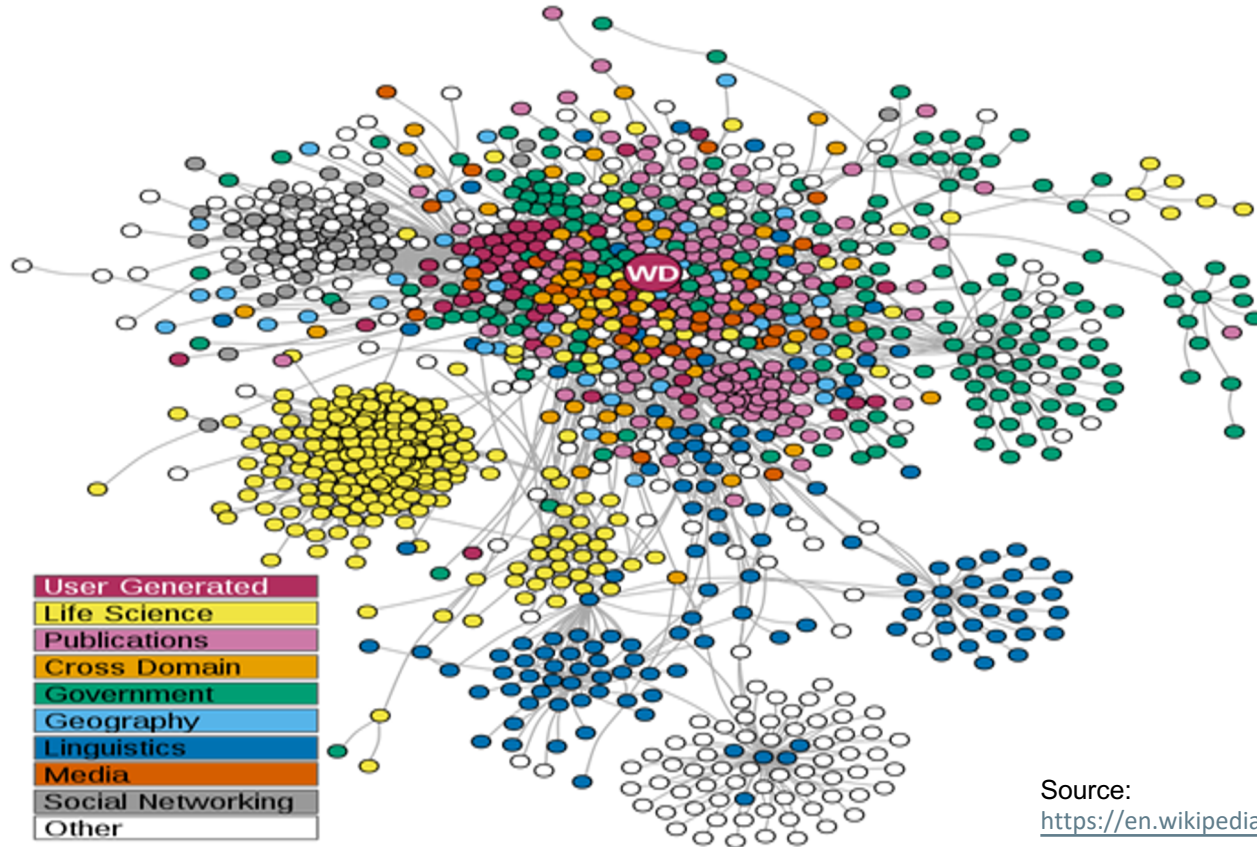
Overview

- Linked Open Data, Wikidata, RDA and BIBFRAME
- WikiProject Chinese Culture and Heritage
- Characteristics of Wikidata on Chinese Related Topics
- Challenges in Creating Chinese Related Wikidata Entries
- Data Model for Chinese Female Poets
- Querying Wikidata
- LOD Education, Collaboration and Applications to Local Libraries
- Lessons Learned and Opportunities
- References & Resources

Linked Open Data (LOD)

	Available on the web (whatever format) <i>but with an open licence, to be Open Data</i>
	Available as machine-readable structured data (e.g. excel instead of image scan of a table)
	as (2) plus non-proprietary format (e.g. CSV instead of excel)
	All the above plus, Use open standards from W3C (RDF and SPARQL) to identify things, so that people can point at your stuff
	All the above, plus: Link your data to other people's data to provide context

Wikidata in the Linked Open Data Cloud



Source:
https://en.wikipedia.org/wiki/Linked_data

RDA Update

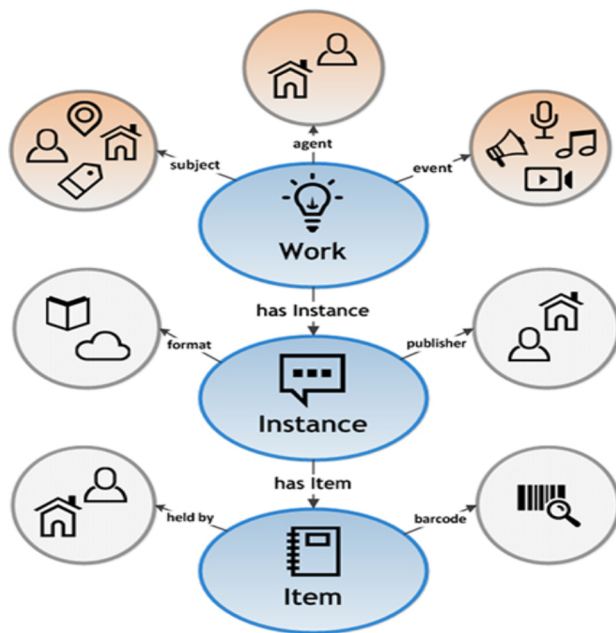
RDA is a package of data elements, guidelines, and instructions for creating library and cultural heritage resource metadata that are well-formed according to international models for user-focused linked data applications.



Source: <http://www.rda-rsc.org/content/about-rda>

BIBFRAME Update

BIBFRAME (Bibliographic Framework) is an initiative to evolve bibliographic description standards to a linked data model, in order to make bibliographic information more useful both within and outside the library community.



Source:

<https://www.loc.gov/bibframe/docs/bibframe2-model.html>

Wikidata: a Platform for Library Linked Open Data

- IFLA Wikidata Working Group (<https://www.ifla.org/node/92837>)
- LD4 Wikidata Affinity Group:
(https://www.wikidata.org/wiki/Wikidata:WikiProject_LD4_Wikidata_Affinity_Group)
- PCC Wikidata Pilot
(https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot)
- LC-BIBFRAME-Wikidata-Project
(https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot/LC-BIBFRAME-Wikidata-Project)

WikiProject Chinese Culture and Heritage (Starting October 2020)

Project coordinator:

Xiaoli Li (University of California, Davis)

Project members:

Sai Deng (University of Central Florida)

Jie Huang (Penn State University)

Jae Jiang (Simmons University)

Jing Jiang (California Digital Library)

Clara Liao (Library of Congress)

Hsianghui Liu-Spencer (Carleton College)

Hong Ma (Loyola University Chicago)

Amanda Xu (National Agricultural Library)

Lihong Zhu (Washington State University)



WikiProject Chinese Culture and Heritage

- **Enhance the existing Wikidata items, such as:**
 - Add Chinese characters
 - Add additional statements
 - Add references
 - Add other IDs (like LC IDs, etc.)
- **Create new Wikidata items**
- **Create data models for special projects**
- **Create new Wikidata properties if needed** (this need to go through Wikidata community approval)

WikiProject Chinese Culture and Heritage

- **Projects Reviewed and Enriched**
 - Eight Great Prose Masters of the Tang and Song
 - Chinese Female Poets
- **Data Models**
 - Data Model for Chinese Female Poets
- **Collaborative Learning**
 - Working with Wikidata
 - SPARQL basics and using the Wikidata Query Service

Characteristics of Wikidata on Chinese Related Topics

- Consider Chinese Language and Culture Characteristics in Creating Wikidata Entries for Chinese Related Topics
- Presentation of Chinese Language
 - Writing System
 - [In more languages](#): English: Su Shi; Chinese: 苏轼; Simplified Chinese: 苏轼; Traditional Chinese: 蘇軾
 - [courtesy name](#): 和仲, [writing system](#): [Chinese characters](#)
 - Pronunciation
 - PinYin, Tones
 - [name in native language](#) : 苏轼 (Chinese (China)) ; [pinyin transliteration](#): sū shì
 - [courtesy name](#): 和仲, [pinyin transliteration](#): Hé Zhòng
 - [Wade-Giles](#), Vernacular spellings of the names: not under various name properties; may input these spellings under the generic [transliteration](#)
- Representation of Chinese Historical and Cultural Uniqueness
 - Time Period, Era Name, Year
 - How to express: 宋仁宗景祐三年十二月十九日 (December 19, the 3rd year of Jingyou, Song Renzong; birth date of Su Shi); date, era name, emperor name
 - [date of birth](#): 8 January 1037
 - [time period](#): [Northern Song Dynasty](#)
 - Era name: [Jingyou Era](#) (constrained to be linked to [Emperor Renzong of Song](#))



Su Shi

Characteristics of Wikidata on Chinese Related Topics

- **Personal Names (courtesy name, art name etc.)**

- From Wikipedia Chinese version: 字子瞻，一字和仲，号东坡居士、铁冠道人
- [courtesy name](#): Zizhan (Chinese: 子瞻), [art-name](#): Dongpo (Chinese: 東坡)

- **Place Names**

- [place of death](#): [Changzhou](#) (also known as: 龍城, 龙城, literally, Dragon city)
- What about other places where Su held official positions, lived, traveled and created poems?
 - Held positions in various places:
[Fengxiang](#), [Hangzhou](#), [Mizhou](#), [Xuzhou](#) (凤翔、杭州、密州、徐州)
 - Properties to consider: [residence](#)
 - Other place related properties: [significant place](#), [location of creation](#)...



Changzhou

Characteristics of Wikidata on Chinese Related Topics

- **Position held (官职)**
 - e.g., 嘉佑六年 (1061) , 大理评事签书凤翔判官 (literally, Dali (Supreme Court) Judge, Clerk (of) Fengxiang (Prefecture) Magistrate) vs. Clerk
- **Pantologist ([occupation](#)):** poet, writer, painter, pharmacist, calligrapher, erudite, historian, politician, musician, essayist
- **Other Considerations: Social, Political, Equity, Diversity & Inclusion (EDI) Complications**
 - How to define Chinese poets?
 - Poets who write in Chinese?
 - Poets who write in Chinese and who were born in mainland China only? Also include poets who were born in Taiwan, Singapore, Malaysia etc.?
 - Poets who write/wrote in Chinese and are/were active in Chinese speaking countries?
 - EDI Considerations
 - [sex or gender](#): female (allow other terms)
 - [ethnic group](#): Han Chinese people
 - [religion](#): Confucianism



Landscape by Zhou Wenjing, featuring part of Liu Zongyuan's poem "Winter Snow" in the upper right corner ("孤舟蓑笠翁，獨釣寒江雪")

Challenges in Creating Chinese Related Wikidata Entries

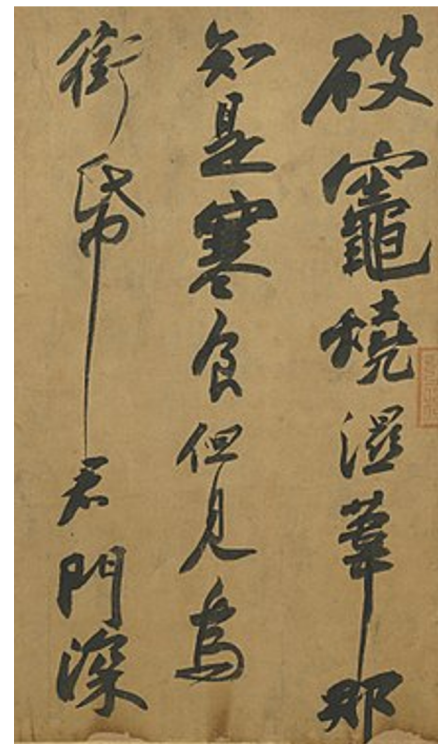
- Describing content via expressing culture uniqueness, or, using a more universal approach?
 - Personal Names
 - [courtesy name \(P1782\)](#) (字)
 - [art-name \(P1787\)](#) (号)
 - [posthumous name \(P1786\)](#) (谥号) v.s.
 - [pseudonym \(P742\)](#)
 - [pen name \(Q127843\)](#)
 - Dynasties, Time Period
 - [country of citizenship \(P27\)](#)
 - data values: Chinese, or Northern Song Dynasty?
 - [time period \(P2348\)](#)
 - data values: year range, dynasty name, or emperor's era name?



Wang Anshi

Challenges in Creating Chinese Related Wikidata Entries

- **Contemporary/ Universal Labels vs. Ancient/ Unique Names**
 - [academic degree \(P512\)](#) (e.g., jinshi 进士, imperial scholar)
 - [field of work \(P101\)](#)
 - [employer \(P108\)](#)
 - [position held \(P39\)](#)
 - [Hanlin Academy Scholar \(Q107382552\)](#) (翰林学士)
 - [Minister of Rites \(Q47175595\)](#) (礼部尚书)
- **Reflecting Historical Changes**
 - Place names
 - [place of birth \(Q1322263\)](#): [Meishan](#) (replaces: [Meishan Prefecture](#))
- **Works, events etc. need to have pre-existing LOD entries**
 - [notable work \(P800\)](#)
 - [significant event \(P793\)](#)



Calligraphy by Su Shi: A detail of The Cold Food Observance (寒食帖)

Data Model for Chinese Female Poets

- **This data model is presented at three levels:**
 - Constant/Const
 - Core
 - Extended/Ext
- **Constant Level Properties**
 - [instance of \(P31\)](#) (Value: human)
 - [sex or gender \(P21\)](#) (Value: female, allow other values)
 - [occupation \(P106\)](#) (Value: poet)
 - [field of work \(P101\)](#) (Value: poetry)

Data Model for Chinese Female Poets

- **Core Level Properties**

- [name in native language \(P1559\)](#)
- [time period \(P2348\)](#)
- [native language \(P103\)](#)
- [languages spoken, written or signed \(P1412\)](#)
- [family name \(P734\)](#)
- [date of birth \(P569\)](#)
- [date of death \(P570\)](#)

- **Extended Level Properties**

- [ethnic group \(P172\)](#)
- [residence \(P551\)](#)



Zhuo Wenjun

Data Model for Chinese Female Poets

- **Extended Level Properties (cont.)**

- [birth name \(P1477\)](#)
- [courtesy name \(P1782\)](#)
- [art-name \(P1787\)](#)
- [pseudonym \(P742\)](#)
- [pen name \(Q127843\)](#)
- [posthumous name \(P1786\)](#)
- [place of birth \(P19\)](#)
- [place of death \(P20\)](#)
- [manner of death \(P1196\)](#)
- [cause of death \(P509\)](#)
- [date of burial or cremation \(P4602\)](#)
- [place of burial \(P119\)](#)
- [father \(P22\)](#)
- [mother \(P25\)](#)
- [sibling \(P3373\)](#)

- [spouse \(P26\)](#)
- [child \(P40\)](#)
- [ancestral home \(P66\)](#)
- [occupation \(P106\)](#)
- [field of work \(P101\)](#)
- [employer \(P108\)](#)
- [part of \(P361\)](#)
- [member of \(P463\)](#)
- [member of political party \(P102\)](#)
- [academic degree \(P512\)](#)
- [educated at \(P69\)](#)
- [award received \(P166\)](#)
- [religion \(P140\)](#)
- [significant event \(P793\)](#)

Data Model for Chinese Female Poets

- **Extended Level Properties**
(cont.)

- [student of \(P1066\)](#)
- [notable work \(P800\)](#)
- [described at URL \(P973\)](#)
- [described by source \(P1343\)](#)
- [Commons Creator page \(P1472\)](#)
- [Commons category \(P373\)](#)
- [topic's main category \(P910\)](#)
- [on focus list of Wikimedia project \(P5008\)](#)
- [copyright status as a creator \(P7763\)](#)
- [has works in the collection \(P6379\)](#)

- [image \(P18\)](#)
- [VIAF ID \(P214\)](#)
- [ISNI \(P213\)](#)
- [Bibliothèque nationale de France ID \(P268\)](#)
- [NCL ID \(P1048\)](#)
- [FAST ID \(P2163\)](#)
- [Academia Sinica authority ID \(P6705\)](#)
- [Union List of Artist Names ID \(P245\)](#)
- [Library of Congress authority ID \(P244\)](#)
- [GND ID \(P227\)](#)
- [WorldCat Identities ID \(P7859\)](#)

Wikidata SPARQL Query for Chinese Female Poets

#find poets; female; writing in Chinese

```
SELECT DISTINCT ?person ?personLabel ?pic ?birthPlace ?birthDate  
WHERE
```

```
{
```

```
?person wdt:P106 wd:Q49757 .      # person whose occupation is poet
```

```
?person wdt:P21 wd:Q6581072 .    # person whose gender is female
```

```
?person wdt:P1412 wd:Q7850 .     # person who writes in Chinese
```

```
?person wdt:P19 ?birthPlace .    # person's birth place
```

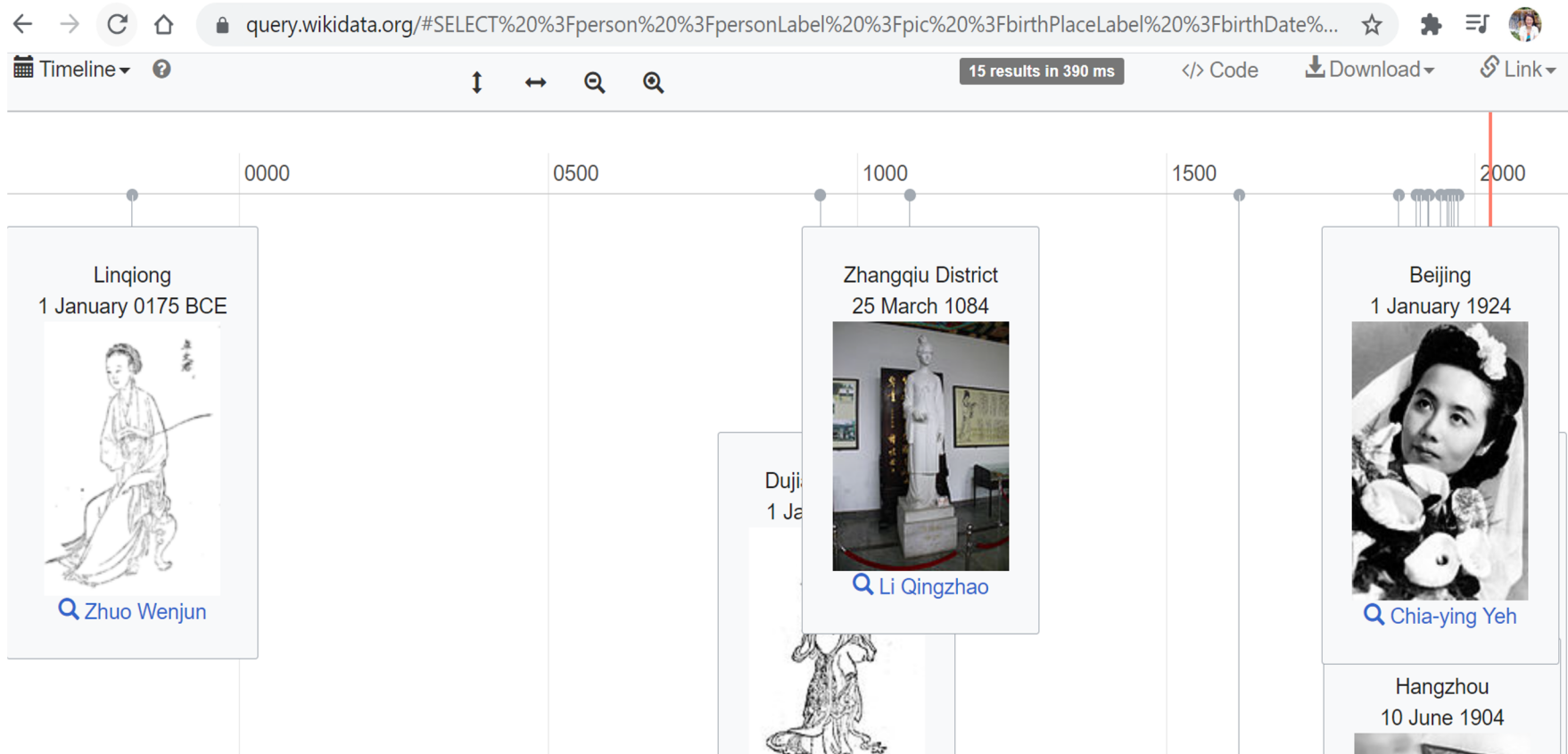
```
?person wdt:P569 ?birthDate .    # person's birth date
```

```
?person wdt:P18 ?pic .           # person's picture
```

```
SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
```

```
}
```

Timeline View of Wikidata for the Chinese Female



Wikidata SPARQL Query for Eight Great Prose Masters of the Tang and Song (Q831237)

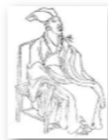
```
SELECT ?person ?personLabel ?pic ?birthPlaceLabel ?coordinates
?birthDate
WHERE
{
  ?person wdt:P19 ?birthPlace .           # person's birth place
  ?birthPlace wdt:P625 ?coordinates .      # GPS coordinates of the birth place
  ?person wdt:P569 ?birthDate .           # person's birth date
  ?person wdt:P18 ?pic .                  # person's picture
  ?person wdt:P361 wd:Q831237 .           # person who is part of Eight Great
  Prose #Masters of Tang and Song Dynasty
  SERVICE wikibase:label { bd:serviceParam wikibase:language "en" }
}
```

Graph View of Wikidata Eight Great Prose Masters of the Tang and Song

query.wikidata.org/#SELECT%20%3Fperson%20%3FpersonLabel%20%3Fpic%20%3FbirthPlaceLabel%20%3Fcoordinate...



Su Xun



Zeng Gong



Su Shi



Ouyang Xiu



Wang Anshi



Han Yu



Su Zhe



Liu Zongyuan

Map View of Wikidata for Eight Great Prose Masters of the Tang and Song

 query.wikidata.org/#SELECT%20%3Fperson%20%3FpersonLabel%20%3Fpic%20%3FbirthPlaceLabel%20%3Fcoordinate...



LOD Education, Collaboration and Applications to Local Libraries

Wikidata Professional Development Training Modules

- [Orientation](#)
- [Introduction to Wikidata](#)
- [Databases and Linked Data](#)
- [The Wikidata community](#)
- [Evaluating Data on Wikidata](#)
- [Adding to Wikidata](#)
- [Querying Wikidata - SPARQL](#)
- [Wikidata WikiProjects](#)

LD4 Conference

- [link](#)

WikiProject LD4 Wikidata Affinity Group Calls

- [Cradle](#)
- [OpenRefine](#)
- [QuickStatements](#)
- [SPARQL query tutorial](#)
- [Wikidata references](#)
- [WikiProjects](#)

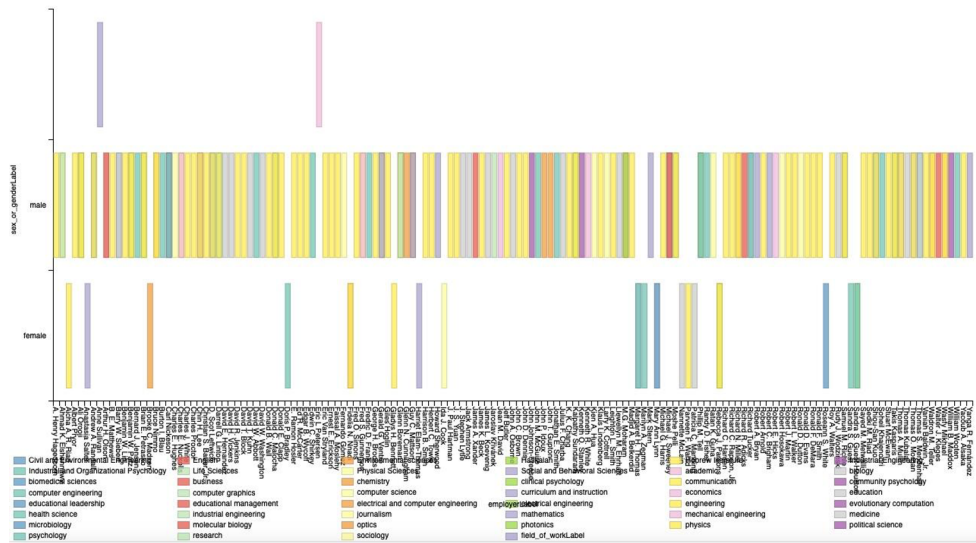
WikiProject PCC Wikidata Pilot Projects

- [link](#)

LOD Education, Collaboration and Applications to Local Libraries

- **Wikidata Projects at Local Libraries**

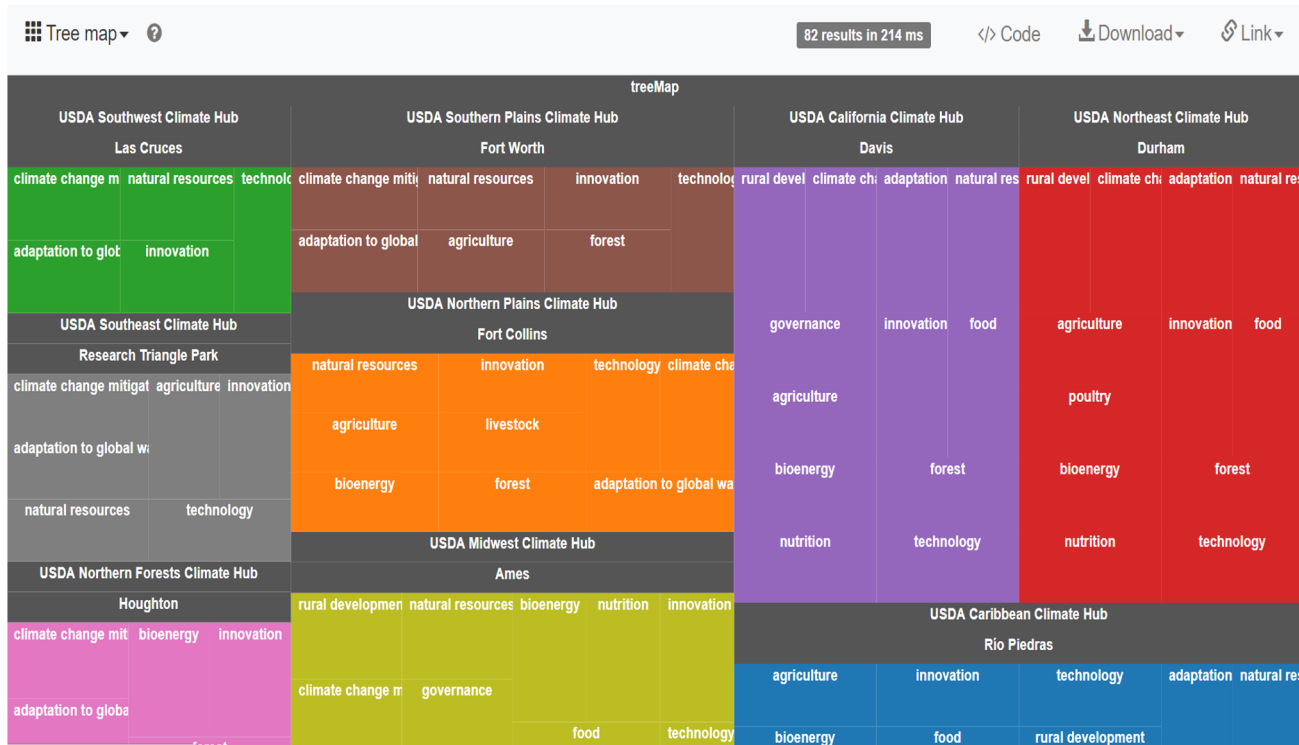
- Pemberton Correspondence Collection
 - Pemberton Family members, e.g.,
 - Phineas Pemberton (Q94253673)
<https://www.wikidata.org/wiki/Q94253673>
- Theses and Dissertation Collection
 - Advisors, e.g.,
 - Richard R. Adicks (Q57704630)
<https://www.wikidata.org/wiki/Q57704630>
- Added Wikidata entries to the UCF Libraries' catalog or IR



Bar Chart: Gender and Discipline Distribution for Advisors

LOD Education, Collaboration and Applications to Local Libraries

- **Wikidata Projects at Local Libraries**
- [WikiProject PCC Wikidata Pilot/National Agricultural Library \(NAL\)](#)
 - Create Wikidata entries for USDA Climate Hubs researchers and corporate bodies
 - Provide the LCCN in Wikidata and add an 024 field to the NACO record



Tree Map: [USDA Climate Hubs](#) (Names, Locations, Field of Work)

Lessons Learned

- It is important to define property level (constant/core/extended), and consistently markup Wikidata items according to data models and source info, e.g. the birth date of Chia-ying Yeh has two values: January 1, 1924 and July 1, 1924
- Need data input standard, e.g. [Alexa Dugan](#) has description “researcher (ORCID 0000-0001-8626-1479)”, but in the publication, he was a Natural Resource Specialist with the U.S. Forest Service. What should be the description?
- May need built-in field-level validation/constraint for the value of a property in a statement, e.g. [item quality evaluator](#) will only give item quality score;
- The key of using query to pull data is to use general conditions/properties or common properties that all target Wikidata have;
- There is a learning curve to use the existing tools for batch-editing and workflow integration, e.g. OpenRefine, QuickStatements, Wikibase.

Opportunities

- Wikidata can lower the barriers for a library to adopt Linked Open Data;
- Wikidata enables creation of unique identifier and identity management on the web;
- Wikidata enables human and machines to work collaboratively to enrich Wikidata items;
- Wikidata powers a lot of other systems, e.g. [Wikidata items are reused widely](#)
 - [Wikidata statements can be referenced to both Wikipedia and other sources](#) (April 12, 2021, 73,178,984 (5.8%) (Wikipedia), 847,353,272 (67.91%) (Other sources));
- Wikidata enables machine processing of authority data, reuse, inference, and distribution;
- Wikidata supports multilingual input and display of Wikidata items;
- Wikidata enables creation of rich data that may be excluded or unavailable in traditional bibliographic data;
- Along with RDA/RDF and BIBFRAME, Wikidata enables global representation and linking
- Enrich bibliographic data in RDA/RDF and BIBFRAME with Wikidata;
- Enable enhanced discovery and representation in knowledge graphs: bring in wikidata into an info box as knowledge panel in discovery interface or search engine interface;
- Enable search support, e.g. type ahead during a search.

References & Resources

Allison-Cassin, Stacy & Scott, Dan (2018, May 4). Wikidata: a platform for your library's linked open data. *Code4lib Journal*, 40. Retrieved from

<https://journal.code4lib.org/articles/13424>

Khan, Huda & Usang, Astrid (2021, June 18). Linked Data + Discovery. *ALA IRRT Webinars*. Retrieved from [https://drive.google.com/file/d/1tx-](https://drive.google.com/file/d/1tx-Frd0RtSXXsWPBYvaOpxgHEkv_9mm3/view)

[Frd0RtSXXsWPBYvaOpxgHEkv_9mm3/view](https://drive.google.com/file/d/1tx-Frd0RtSXXsWPBYvaOpxgHEkv_9mm3/view)

Klein, M., & Kyrios, A. (2013, October 14). VIAFbot and the Integration of Library Data on Wikipedia. *Code4Lib Journal*, 22. Retrieved from <http://journal.code4lib.org/articles/8964>

Linked Open Data Cloud. Retrieved from <https://lod-cloud.net/>

Riemer, John (2021, May 7). Wikidata Pilot. *PCC Operations Committee Meeting*.

References & Resources

OCLC Research. (2018). International Linked Data survey. Retrieved from <https://www.oclc.org/research/themes/data-science/linkedata/linked-data-survey.html>

Wallis, Richard (2021, Feb. 11) . Follow me to the Library: Bibliographic Data in a Discovery Driven World. *International Conference Bibliographic Control in the Digital Ecosystem 2021*.

Wikidata Item Quality Evaluator. Retrieved July 3, 2021 from <https://item-quality-evaluator.toolforge.org/>

Wikidata: WikiProject PCC Wikidata Pilot. Retrieved July 3, 2021 from https://www.wikidata.org/wiki/Wikidata:WikiProject_PCC_Wikidata_Pilot

Wikidata Stats. Retrieved July 3, 2021 from <https://wikidata-todo.toolforge.org/stats.php>

Image Sources

Berners-Lee, T. (2006). Linked Open Data 5 star. <https://www.w3.org/DesignIssues/LinkedData.html>

BIBFRAME 2.0 Model.

<https://www.loc.gov/bibframe/docs/bibframe2-model.html>

Calligraphy by Su Shi: A detail of The Cold Food Observance (寒食帖). https://en.wikipedia.org/wiki/Su_Shi

Changzhou. <https://www.wikidata.org/wiki/Q57970>

Landscape by Zhou Wenjing. https://en.wikipedia.org/wiki/Liu_Zongyuan

Su Shi (image). <https://www.wikidata.org/wiki/Q36020>

Wang Anshi (image). <https://www.wikidata.org/wiki/Q319618>

Wikidata in the Linked Open Data Cloud. https://en.wikipedia.org/wiki/Linked_data

Xun, Jian. Li Qingzhao (Cover Image). https://en.wikipedia.org/wiki/Li_Qingzhao

Zeng, Gong. Letter to a friend. https://en.wikipedia.org/wiki/Zeng_Gong

Zhuo Wenjun. <https://www.wikidata.org/wiki/Q3358500>

Contact & Questions

Amanda Xu

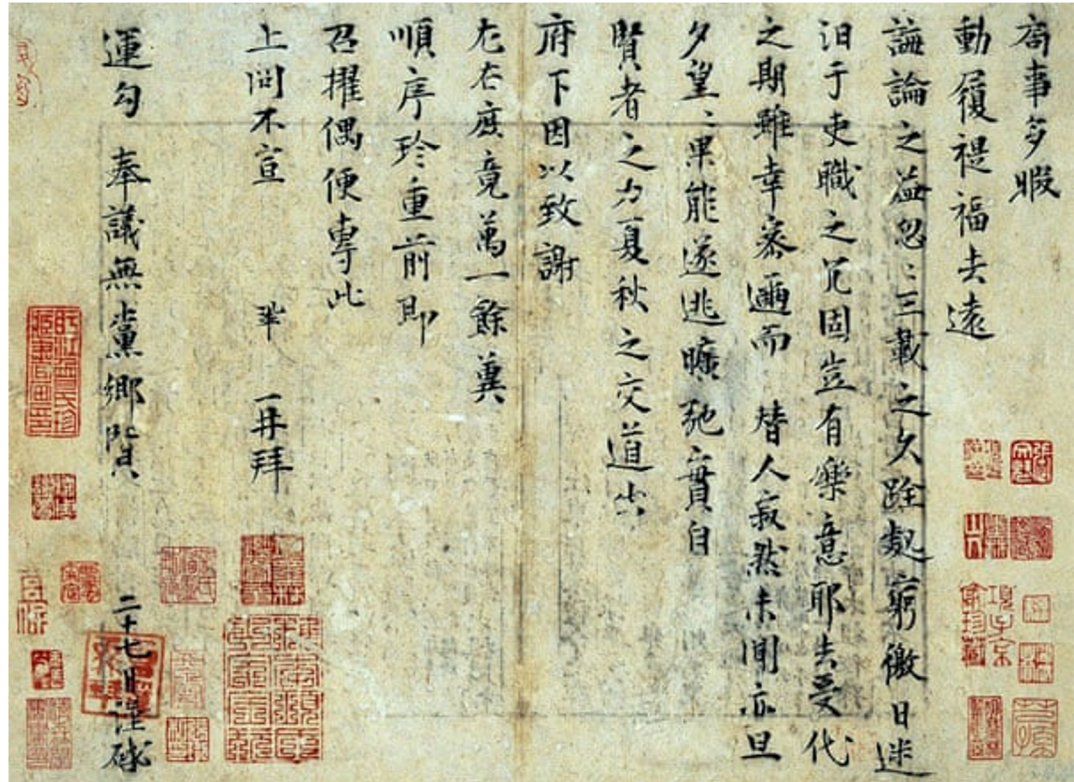
amanda.xu@usda.gov

Lihong Zhu

lzhu2@wsu.edu

Sai Deng

sai.deng@ucf.edu



Letter to a friend written by Zeng Gong (Source: https://en.wikipedia.org/wiki/Zeng_Gong)